Recorded 1/9/07 187

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.6

Run on: November 17, 2003, 09:13:37 ; Search time 3691 Seconds
                                              (without alignments)
                                          10496.184 Million cell updates/sec.

Title: US-10-043-539A-1
Perfect score: 947
Sequence: 1 gtttcaaatcggtggagg.........gtttcataaactgaatcgat 947

Scoring table: IDENTITY NUC
                Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl.*
            1: gb_ba:*
            2: gb_htg:*
            3: gb_in:*
            4: gb_om:*
            5: gb_ov:*
            6: gb_pat:*
            7: gb_ph:*
            8: gb_pl:*
            9: gb_p:*
           10: gb_ro:*
           11: gb_sts:*
           12: gb_sy:*
           13: gb_un:*
           14: gb_vi:*
           15: gb_ov:*
           16: em_fun:*
           17: em_hum:*
           18: em_in:*
           19: em_mu:*
           20: em_om:*
           21: em_or:*
           22: em_ov:*
           23: em_pat:*
           24: em_ph:*
           25: em_pl:*
           26: em_ro:*
           27: em_sts:*
           28: em_un:*
           29: em_vi:*
           30: em_htg_hum:*
           31: em_htg_inv:*
           32: em_htg_other:*
           33: em_htg_mus:*
           34: em_htg_pln:*
           35: em_htg_rod:*
           36: em_htg_mam:*
           37: em_htg_vrt:*
           38: em_sy:*
           39: em_hcgo_hum:*
           40: em_hcgo_mus:*
           41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a

SUMMARIES

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	947	100.0	947	1	AP207701	Staphyloc
2	826.4	87.3	304050	1	AP004829	Staphyloc
3	826.4	87.3	307750	1	AP003136	Staphyloc
4	824.8	87.1	348650	1	AP003364	Staphyloc
5	369.8	39.0	384	6	AX620460	Sequence
6	307	32.4	300698	1	AE016750	Staphyloc
7	199.4	21.1	834	6	AX620416	Sequence
8	96.2	10.2	8056	6	AX599046	Sequence
9	92.8	9.8	170627	2	AC125567	Rattus no
10	91.4	9.7	67970	3	PFMAL1P3	Plasmodiu
11	90.6	9.6	14422	3	AF466146	Melipona
12	90.4	9.5	86826	3	AL034556	Plasmodiu
13	90	9.5	8056	3	AX599046	Sequence
14	87.6	9.3	258658	3	AE014832	Plasmodiu
15	87.2	9.2	8056	6	AX598900	Sequence
16	85.2	9.0	286208	2	AC117140	Rattus no
17	85	9.0	840	8	CNS0180K	Botrytis
18	84.8	9.0	2009	9	AX457067	Sequence
19	84.6	8.9	250029	3	AE014820	Plasmodiu
20	84.2	8.9	259474	9	HUAC004605	Homo sapi
21	83.6	8.8	175544	2	AC117342	Rattus no
22	83	8.8	137411	9	AC092066	Homo sapi
23	82.8	8.7	1090	3	AB084761	Drosophil
24	82.8	8.7	169046	5	AL928828	Zebrafish
25	82.4	8.7	348600	1	AB063521	Wiggleswo
26	82.4	8.7	349980	6	AX344555	Sequence
27	82.2	8.7	196875	2	AC098095	Rattus no
28	82	8.7	6292	9	AX251493	Sequence
29	82	8.7	233877	9	AC093798	Homo sapi
30	81.6	8.6	349751	3	PFMAL4P3	Plasmodiu
31	81.4	8.6	156060	2	AC004153	Plasmodiu
32	81.4	8.6	250707	3	AE014848	Plasmodiu
33	81.4	8.6	347050	2	PFA929351	Plasmodiu
34	81.2	8.6	99003	2	AL390756	Homo sapi
35	81.2	8.6	110000	2	PFMAL13_09	Continuation (10 o
36	81.2	8.6	318221	2	PFMAL13F3	Plasmodiu
37	81	8.6	162445	9	AL158151	Human DNA
38	81	8.6	164640	2	BX005461	Danio rer
39	81	8.6	168558	2	BX322549	Danio rer
40	80.8	8.5	84563	5	AC096885	Danio rer
41	80.6	8.5	313050	3	PFA929352	Plasmodiu
42	80.2	8.5	8056	6	AX598900	Sequence
43	80.2	8.5	60604	2	AC023466	Homo sapi
44	80.2	8.5	181792	9	AC098822	Homo sapi
45	80	8.4	112695	8	AC119418	Medicago

Description (column):
AP207701 Staphyloc
AP004829 Staphyloc
AP003136 Staphyloc
AP003364 Staphyloc
AX620460 Sequence
AE016750 Staphyloc
AX620416 Sequence
AX599046 Sequence
AC125567 Rattus no
PFMAL1P3 Plasmodiu
AF466146 Melipona
AL034556 Plasmodiu
AX599046 Sequence
AE014832 Plasmodiu
AX598900 Sequence
AC117140 Rattus no
CNS0180K Botrytis
AX457067 Sequence
AE014820 Plasmodiu
AC004605 Homo sapi
AC117342 Rattus no
AC092066 Homo sapi
AB084761 Drosophil
AL928828 Zebrafish
AB063521 Wiggleswo
AX344555 Sequence
AC098095 Rattus no
AX251493 Sequence
AC093798 Homo sapi
AL035476 Plasmodiu
AC004153 Plasmodiu
AE014848 Plasmodiu
AL929351 Plasmodiu
AL390756 Homo sapi
AL049184 Plasmodiu
AL158151 Human DNA
BX005461 Danio rer
BX322549 Danio rer
AC096885 Danio rer
AL929352 Plasmodiu
AX598900 Sequence
AC023466 Homo sapi
AC098822 Homo sapi
AC119418 Medicago

ALIGNMENTS

RESULT 1
AF207701
LOCUS   AF207701   947 bp   DNA   linear   BCT 01-DEC-2000
DEFINITION   Staphylococcus aureus putative transposase gene, partial cds; and
             SarR (sarR) gene, complete cds.
ACCESSION   AF207701
VERSION   AF207701.1 GI:11493941
KEYWORDS   .
SOURCE   Staphylococcus aureus
  ORGANISM   Staphylococcus aureus
             Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1 (bases 1 to 947)
  AUTHORS   Cheung,A.L. and Manna,A.C.
  TITLE   Characterization of SarR, a modulator of sar expression in
             Staphylococcus aureus

```
JOURNAL    Unpublished
REFERENCE  2  (bases 1 to 947)
  AUTHORS   Cheung,A.L. and Manna,A.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-NOV-1999) Microbiology, Dartmouth Medical School,
            College St, Vail 206, Hanover, NH 03755, USA
FEATURES    Location/Qualifiers
  source    1..947
            /organism="Staphylococcus aureus"
            /mol_type="genomic DNA"
            /strain="RN6390"
            /db_xref="taxon:1280"
  CDS       <1..70
            /codon_start=2
            /transl_table=11
            /product="putative transposase"
            /protein_id="AAG35716.1"
            /db_xref="GI:11493943"
            /translation="FSKSVEVHEKVIGHPLKIKKYQ"
  gene      208..555
            /gene="sarR"
            /note="regulatory gene"
  CDS       208..555
            /gene="sarR"
            /note="similar to SarA"
            /codon_start=1
            /transl_table=11
            /product="SarR"
            /protein_id="AAG35715.1"
            /db_xref="GI:11493942"
            /translation="MSKINDINDLVNATFGVKFFRDTKKKFNLNYESIYILNHILRS
            ESNEISSKEIAKCSEFKPYLTKALQKLKDLAGLLSKKRSLQDERTVIVVTDTQKANI
            QKLISELEEYIKN"
BASE COUNT     374 a    112 c    145 g    316 t
ORIGIN

Query Match       100.0%;  Score 947;  DB 1;  Length 947;
Best Local Similarity  100.0%;  Pred. No. 7.7e-126;
Matches 947;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 GTTTCAAAATCGGTGGAGGTGCATGAAAAGTTATTGGGCATTTTGAAAATAAAAAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GTTTCAAAATCGGTGGAGGTGCATGAAAAGTTATTGGGCATTTTGAAAATAAAAAA 60

Qy    61 ATATCAATAAGTTGGAGTCATTACCGAATTTTATACTTATTGTTTAGAATGAACTTTA 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ATATCAATAAGTTGGAGTCATTACCGAATTTTATACTTATTGTTTAGAATGAACTTTA 120

Qy   121 TAACATAGTTGGATAGAGTTTTCGAATTTAATACATTAAATGTGAACCTGCTACAACAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TAACATAGTTGGATAGAGTTTTCGAATTTAATACATTAAATGTGAACCTGCTACAACAG 180

Qy   181 ATGTGCATCAGAAGGAGTGGTTTAATAATGAGTAAAATTAATGATATTAATGATTAGTC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ATGTGCATCAGAAGGAGTGGTTTAATAATGAGTAAAATTAATGATATTAATGATTAGTC 240

Qy   241 AACGCAACATTCAAGTTAAGAAGTTTTTCAGAGATACAAAAAGAAGTTCAATTGAAC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AACGCAACATTCAAGTTAAGAAGTTTTTCAGAGATACAAAAAGAAGTTCAATTGAAC 300

Qy   301 TATGAAGAAATTTATATTTTAAATCATATTTTAAGGAAGTGAGTCTAACGGAAATCTCATCT 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TATGAAGAAATTTATATTTTAAATCATATTTTAAGGAAGTGAGTCTAACGGAAATCTCATCT 360

Qy   361 AAAGAGATTGCTAAGTGCTCAGAGTTCAAACCTACTATTTAACTAAAGCTTTACAAAAG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGAGATTGCTAAGTGCTCAGAGTTCAAACCTACTATTTAACTAAAGCTTTACAAAAG 420

Qy   421 CTAAAAGATTTAAAATTGTTATCAAAGAAAAGAAGTTACAAGACGAAAGAACAGTTATT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTAAAAGATTTAAAATTGTTATCAAAGAAAAGAAGTTACAAGACGAAAGAACAGTTATT 480

Qy   481 GTTTATGTTACAGATCACACAAAAAGCAAATATTCAAAAACTGATTTCAGATTAGAAGAA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GTTTATGTTACAGATCACACAAAAAGCAAATATTCAAAAACTGATTTCAGATTAGAAGAA 540

Qy   541 TACATTAAAAATTAAATCAAGGTTAATGGCGTTAATACATTGAACGATAACAATTTAT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TACATTAAAAATTAAATCAAGGTTAATGGCGTTAATACATTGAACGATAACAATTTAT 600

Qy   601 TAATACGAAGTTATTATTCAGCATGGGACATAAAAATTAACTAAAATTTAAATATTGA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TAATACGAAGTTATTATTCAGCATGGGACATAAAAATTAACTAAAATTTAAATATTGA 660

Qy   661 AGATGCTTTAATTAAAGTTAAAGACCAGCCATACCTTATTTCAGCTTATTAAGCTTGACA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AGATGCTTTAATTAAAGTTAAAGACCAGCCATACCTTATTTCAGCTTATTAAGCTTGACA 720

Qy   721 CAAGGTACACTAGTCTTTTATTTAATATTTCTTAGAAAATCAAGTTACGATCATAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CAAGGTACACTAGTCTTTTATTTAATATTTCTTAGAAAATCAAGTTACGATCATAA 780

Qy   781 ATATTTTCTGCGATATAGCTTTGGATGGTTCCAAGTATTTCTCTATAATTTGTGTGCGA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATATTTTCTGCGATATAGCTTTGGATGGTTCCAAGTATTTCTCTATAATTTGTGTGCGA 840

Qy   841 TAAGCAAAAATTCTAACTGCAAAACCATGTGTAGGCAATTGAGAAATAGCAACACCACAA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TAAGCAAAAATTCTAACTGCAAAACCATGTGTAGGCAATTGAGAAATAGCAACACCACAA 900

Qy   901 TCGGATGTATTGCTATAAGAACTAATGGTTTCATAAACTGAATCGAT 947
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCGGATGTATTGCTATAAGAACTAATGGTTTCATAAACTGAATCGAT 947
```

```
                Copyright  GenCore version 5.1.6
                Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         November 17, 2003, 09:11:34 ; Search time 316 Seconds
                                            (without alignments)
                                            8089.772 Million cell updates/sec

Title:          US-10-043-539A-1
Perfect score:  947
Sequence:       1 gtttccaaatcggtggagg.........gtttcataaactgaatcgat 947

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       252756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters:      5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      N_Geneseq_19Jun03;*
        1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
        2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
        3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
        4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
        5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
        6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
        7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
        8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
        9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
       10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
       11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
       12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
       13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
       14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
       15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
       16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
       17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
       18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
       19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
       20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
       21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
       22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
       23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
       24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
       25:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Match Length	DB	ID	Description
1	329	34.7	400	18	AAV75836	Staphylococcus aur
2	261.4	27.6	387	24	ABN92693	Staphylococcus epi
3	96.2	10.2	8056	25	AB210246	Haematopoietic cel
4	90	9.5	8056	25	AB210246	Haematopoietic cel
5	87.2	9.2	8056	25	AB210100	Haematopoietic cel
6	82	8.7	6292	22	AAS46735	Tumour suppressor
7	80.2	8.5	8056	25	AB210100	Haematopoietic cel
8	79.4	8.4	101	18	AAV75910	Staphylococcus aur
c 9	77.6	8.2	867	24	ABN92692	Staphylococcus epi
c 10	75.6	8.0	529	21	AAA66121	E. coli proliferat
c 11	75.2	7.9	12079	24	ABQ76793	pUC19 promoter-ter
c 12	75.2	7.9	12079	24	ABV74271	Plant specific exp
c 13	75.2	7.9	12085	24	ABQ76792	pUC19 promoter-ter
c 14	75.2	7.9	12085	24	ABV74270	Plant specific exp
c 15	75.2	7.9	12093	24	ABQ76791	pUC19 promoter-ter
c 16	75.2	7.9	12093	24	ABV74269	Plant specific exp
c 17	75.2	7.9	12241	24	ABQ73049	Tomato anthocyanin
c 18	75.2	7.9	12241	24	AAD36732	Binary vector DNA
c 19	75.2	7.9	13002	24	ABQ76794	pUC19 promoter-ter
c 20	75.2	7.9	13002	24	ABV74272	Plant specific exp
c 21	75.2	7.9	13737	21	AAA54212	Transformation vec
c 22	75.2	7.9	13905	24	ABQ76795	pUC19 promoter-ter
c 23	75.2	7.9	13905	24	ABV74273	Plant specific exp
c 24	75.2	7.9	14446	24	AAS17548	Plasmid pTM036. S
c 25	75.2	7.9	15430	24	ABQ76796	pB-DHGLA encoding
c 26	75.2	7.9	15430	24	ABV74274	Plant specific exp
c 27	75.2	7.9	17752	24	ABQ76797	pBARA1 encoding de
c 28	75.2	7.9	17752	24	ABV74275	Plant specific exp
c 29	75.2	7.9	19705	24	ABA92074	Vector plasmid pBT
c 30	75.2	7.9	20119	22	AAH26041	Plasmid pNOV524 en
c 31	75.2	7.9	20174	24	ABA92073	Transformation vec
c 32	75.2	7.9	26019	22	AAH26040	Plasmid pNOV523 en
c 33	75	7.9	6664	24	AAS61368	Human gene regulat
c 34	74.6	7.9	4086	22	AAH54401	S. epidermidis gen
c 35	74.4	7.9	778	24	ABQ15588	Oligonucleotide fo
c 36	74.4	7.9	778	24	ABQ15589	Oligonucleotide fo
c 37	73.2	7.7	6191	24	ABN80161	Human chemically m
c 38	73.2	7.7	6191	24	ABL70282	Chemically treated
c 39	73.2	7.7	6191	24	ABL33217	Human immune syste
c 40	73.2	7.7	6191	24	ABK31307	Signal transductio
c 41	73.2	7.7	15548	24	ABL34155	Human immune syste
c 42	72.2	7.6	3967	22	AAH54106	S. epidermidis gen
c 43	72	7.6	1501	25	AB210188	Haematopoietic cel
c 44	72	7.6	2145	21	AAZ56364	Escherichia coli f
c 45	72	7.6	4142	22	AAH54879	S. epidermidis gen

ALIGNMENTS

```
RESULT 1
AAV75836
ID   AAV75836 standard; DNA; 400 BP.
XX
AC   AAV75836;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #1525.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key            Location/Qualifiers
FH
FT   misc_feature   301..360
FT                  /tag= a
FT                  /note= "these bases represent a line of missing text in
FT                  the sequence listing in the specification. They
FT                  are included to maintain the nucleotide numbering
FT                  given in the specification for this DNA sequence"
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;  97EP-0100117.
XX
```

PR   05-JAN-1996;   96US-0009861.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
PI   Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT   stored on computer readable medium and used in the production of
PT   anti-S.aureus vaccines
XX
PS   Claim 1; Page 1984-1985; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the
CC   computer readable medium.
XX
SQ   Sequence 400 BP; 155 A; 36 C; 44 G; 104 T; 61 other;

Query Match   34.7%;   Score 329;   DB 18;   Length 400;
Best Local Similarity   84.4%;   Pred. No. 2.2e-46;
Matches 329;   Conservative   0;   Mismatches   61;   Indels   0;   Gaps   0;

Qy   274 GATACAAAAAGAAGTCAATTGAACTATGAAGAAATTTATATTTTAAATCATATTTTA 333
                |||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   1 GATACAAAAAGAAGTCAATTGAACTATGAAGAAATTTATATTTAAATCATATTTTA 60

Qy   334 AGAAGTGAGTCTAACGAAATCTCATCTAAAGAGATGCTAAGTGCTCAGAGTTCAAACCT 393
                ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   61 AGAAGTGAGTCTAACGNAATCTCATCTAAAGAGATGCTAAGTGCTCAGAGTTCAAACCT 120

Qy   394 TACTATTTAACTAAAGCTTTACAAAAGCTAAAAGATTTAAAATTGTTATCAAGAAAAGA 453
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TACTATTTAACTAAAGCTTTACAAAAGCTAAAAGATTTAAAATTGTTATCAAGAAAAGA 180

Qy   454 AGTTTACAAGACGAAAGAACAGTTATTGTTTATGTTACAGATACACAAAAAGCAAATATT 513
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AGTTTACAAGACGAAAGAACAGTTATTGTTTATGTTACAGATACACAAAAAGCAAATATT 240

Qy   514 CAAAAACTGATTCAGAATTAGAAGAATACATTAAAAATTAAAATCAAGGTTAATTGCGTT 573
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAAAAACTGATTCAGAATTAGAAGAATACATTAAAAATTAAAATCAAGGTTAATTGCGTT 300

Qy   574 TAATAACATTGAACGATAACAATTTATTAATACGAAGTTATTATTCAGCATTGGGACAT 633
                |||||||||||||||||||||||||||||||
Db   301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360

Qy   634 AAAATTAACTTAAAATTTAAATATTGAAGA 663
                |||||||||||||||||||||||||||
Db   361 AAAATTAACTTAAAATTTAAATATTGAAGA 390

RESULT 2
ABN92693
ID   ABN92693 standard; DNA; 387 BP.
XX

AC   ABN92693;
XX
DT   24-JUL-2002 (first entry)
XX
DE   Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:2156.
XX
KW   Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW   antibacterial; gene therapy; gene; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   US6380370-B1.
XX
PD   30-APR-2002.
XX
PF   13-AUG-1998;   98US-0134001.
XX
PR   14-AUG-1997;   97US-055779P.
PR   08-NOV-1997;   97US-0649642.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Doucette-Stamm LA,   Bush D;
XX
DR   WPI; 2002-381255/41.
DR   P-PSDB; ABP40148.
XX
PT   Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT   polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS   Disclosure; SEQ ID 2156; 267pp; English.
XX
CC   ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC   frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC   given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC   antibacterial activity and can be used in gene therapy. The sequences
CC   can also be used in the diagnosis and treatment of bacterial infections,
CC   particularly S. epidermidis infections. The sequences can be used to
CC   screen for compounds able to interfere with the S. epidermidis life
CC   cycle or inhibit S. epidermidis infection.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   USPTO web site.
XX
SQ   Sequence 387 BP; 175 A; 39 C; 55 G; 118 T; 0 other;

Query Match   27.6%;   Score 261.4;   DB 24;   Length 387;
Best Local Similarity   81.4%;   Pred. NO. 4.2e-35;
Matches 315;   Conservative   0;   Mismatches   71;   Indels   1;   Gaps   1;

Qy   165 ACCTTGCTACAACAAGATGTGCATCAGAAGGAGTGGTTTAATAATGAGTAAAATTAATGA 224
                ||||||||||||||||||||||| |||||||| ||||||| | |||||||||||||||
Db   1 ACCTTGCTACAACAACAAGATGTGCATCAG-AGGAGTGGTTTATAATGGGAAAAATTAAAGA 59

Qy   225 TATTAATGATTAGTCAACGCAACATTCAAGTTAAGAAGTTTTTCAGATACAAAAAA 284
                |||||||||||||| || || || |||||||| |||||||| |||||||||||||
Db   60 CATCAAGATTTGGTTAATGCGACATTTCAAGTAAAAAAAATTTTTAGAGATACTAAAAA 119

Qy   285 GAAGTTCAATTGAACTATGAAGAAAATTTATATTTTAAGAAGTGAGTC 344
                | ||||||||||||||||||||| |||||||| |||||||||||||
Db   120 GCAATATATATTTAAATATGAAGAAAATTTATTCTTAATCCATTTTGAAAAGTGAATC 179

Qy   345 TAACCAAATCTCATCTAAAGAGATTGCTAAGTGCTCAGAGTTCAAACCTTACTATTTAAC 404
                |||||||||||||||||||||||||||||||||||||||||| ||||||||| |||||||
Db   180 TAACGAAATATCTTCAAAAGAAAATTGCTACACATGTTCAGAGTTTAAACCGTATTATTTAAC 239

Qy   405 TAAAGCTTTACAAAAGCTAAAAGCTAAAAGATTTAAAATTGTTATCAAGAAAAGAAGTTTACAGA 464
                |||||| || || ||||||| ||||| ||||||||| ||||| |||||| |||||||||||||
Db   240 TAAAGCATTACAAAAATTAAAAGAGATTTAAAATTAAATTTACTATCTAAAAAGCAAATATTCAAAAACTGAT 299

Qy   465 CGAAAGAACAGTTATTGTTTATGTTACAGATACACAAAAAGCAAATATTCAAAAACTGAT 524
                || |||||||||||||| ||| || || |||| |||||| ||| ||| | |||||||
Db   300 TGAAAGAACAGTAATTGTTATTGTATCAGATGACAACGAGAAAAAATAAAAAAAATTGAT 359